

# PROBLEMAS ESTADÍSTICA I

## INGENIERÍA TÉCNICA EN INFORMÁTICA

### CURSO 2002/2003

## Estadística Descriptiva Unidimensional

1. Un edificio tiene 45 apartamentos con el siguiente número de inquilinos:

2	1	3	5	2	2	2	1	4	2	6	2	4	3	1
2	4	3	1	4	4	2	4	4	2	2	3	1	4	2
3	1	5	2	4	1	3	2	4	4	2	5	1	3	4

Se pide

- Construir la tabla de frecuencias de la variable número de inquilinos.
  - Dibujar el polígono de frecuencias y el diagrama de frecuencias relativa acumulado.
  - Calcular media, mediana y moda del número de inquilinos.
  - ¿Cuál es el porcentaje de apartamentos con más de cuatro inquilinos?, ¿Y con menos de 3?.
2. Snyder y Chrissis (1990) presentaron un algoritmo híbrido para resolver un problema de Programación Matemática Polinomial Cero-Uno. El algoritmo incorpora una combinación de conceptos pseudobooleanos y procedimientos de enumeración implícitos probados y comprobados. Se resolvieron 52 problemas al azar utilizando el algoritmo híbrido, los tiempos de resolución (tiempo de CPU en segundos) se listan en la siguiente tabla.

0.045	1.045	0.136	0.894	0.379	0.136
0.336	0.258	0.070	0.506	0.088	0.242
0.639	0.912	0.412	0.361	0.788	0.579
0.267	0.567	0.182	0.036	0.394	0.209
0.445	0.179	0.118	0.333	0.554	0.258
0.182	0.070	0.985	0.670	0.888	0.136
0.091	0.600	0.291	0.327	0.130	0.145
0.170	0.227	0.064	0.194	0.209	0.258
0.046	0.045	0.049	0.079		

Se pide:

- Agrupar en clases de igual amplitud comenzando en 0 y finalizando en 1.05, y construir la tabla de frecuencias de los tiempos de resolución de los problemas.
- Representar el histograma frecuencias y el polígono de frecuencias acumuladas.
- Calcular las principales medidas de tendencia central (media, mediana y moda).
- Calcular las medidas de dispersión: varianza y coeficiente de variación de Pearson. ¿Es la media representativa de los datos?
- Determinar el recorrido intercuartílico.
- ¿Qué podemos decir de la simetría de la variable?.

3. Una compañía eléctrica ha realizado una investigación estadística en relación con las causas que han provocado la interrupción del suministro de energía eléctrica en la red que abastece y se han obtenido los siguientes datos.

NA	NA	AT	AT	NA	FF	NA	OC	FF	FF
NA	NA	AT	NA	NA	NA	AT	AT	AT	AT
OC	AE	AT	AT	AT	AT	AT	AT	NA	NA

donde las abreviaturas utilizadas significan:

- FF: Fallo Elemento.
- AT: Atmosféricas.
- AE: Agentes Externos.
- FM: Falsas Maniobras.
- AP: Mala Actuación de Protecciones.
- OC: Otras Causas.
- NA: No aclaradas.

Sin organizar este conjunto de datos no podemos responder de manera inmediata a preguntas del tipo:

- (a) ¿Más de la mitad de interrupciones se debieron a causas atmosféricas?.
  - (b) ¿El porcentaje de interrupciones no aclaradas es superior al debido a causas atmosféricas?.
  - (c) Construir la tabla de frecuencias de la variable estadística en estudio. ¿De qué tipo se trata?.
  - (d) Representar un diagrama de rectángulos y uno circular.
4. Los datos que se muestran a continuación son el resultado de la duración de bombillas de 60 Watios, expresadas en horas que una bombilla ha lucido antes de fundirse.

963.4	1132.5	1021.7	1087.3	950.1	910.0	1045.0	1114.1	1085.1	1002.4
1175.9	1052.1	976.1	1114.1	932.9	1060.1	1123.6	1130.0	951.9	1187.0
1074.7	1092.2	987.2	1089.5	810.1	1066.2	1121.9	1121.9	1001.2	1023.8
887.1	874.8	1109.4	1133.5	1042.1	900.0	962.1	1058.4	956.5	909.2
1078.3	1001.7	1099.3	1072.8	1129.8	1142.2	939.0	1149.9	1089.3	1059.2
1065.4	1198.2	1075.6	866.0	995.3	1152.8	1092.7	1028.1	1033.2	1078.7
1124.8	1003.4	901.3	922.0	1129.8	1093.1	1132.7	822.9	949.9	1092.1
1083.8	1130.2	988.8	1121.8	1021.3	1049.9	1083.8	1114.4	1043.7	1066.3
882.6	1092.7	1074.1	1142.0	1102.1	1078.4	989.8	1166.2	1023.9	984.3
1055.5	1143.8	1109.8	822.5	970.0	1021.3	1021.0	1122.8	1056.2	1112.1

Se pide:

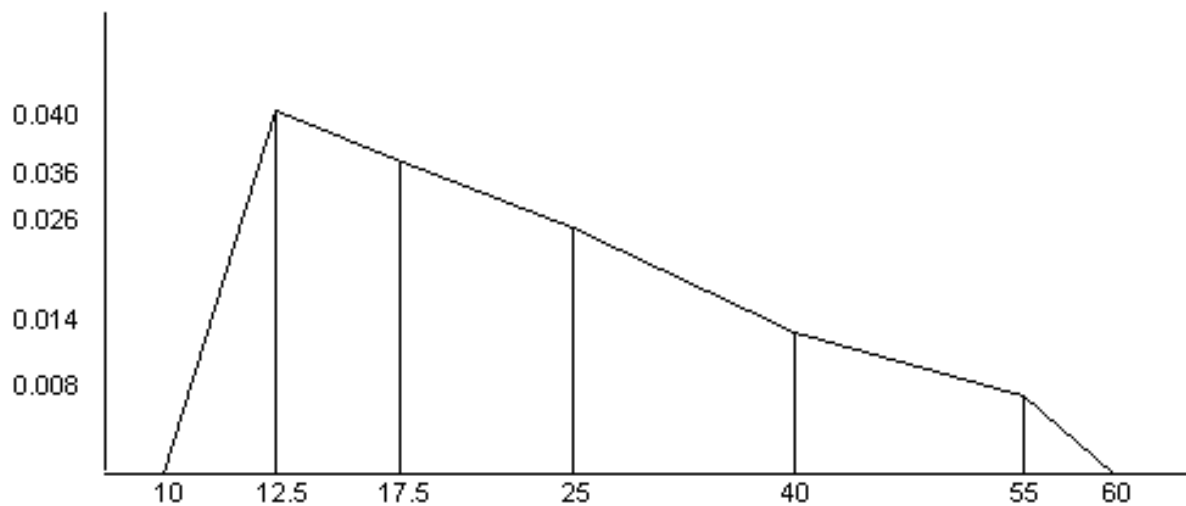
- (a) Agrupar los datos en clases de amplitud constante, empezando en 800 y finalizando en 1200; y construir la distribución de frecuencias del tiempo de duración de las bombillas.
- (b) Representar el hitograma y el polígono de frecuencias para la variable en estudio.
- (c) Hallar la media, mediana y moda.
- (d) Calcular la varianza y los cuartiles.
- (e) ¿Cuales son los valores de la variable que dejan el 30% de las observaciones centrales?.

5. Una empresa está realizando un estudio sobre la capacidad de ahorro de los individuos de una determinada región. Entre otros datos, obtuvo la distribución del ahorro anual, habiéndose construido una tabla completa que posteriormente se extravió, quedando únicamente una fotocopia en mal estado con muchos datos borrosos. Reconstruya la tabla original a partir de la siguiente incompleta. Los datos aparecen en miles de pesetas.

Intervalo	$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	Amplitud
20-50	35	2	2			
- 60			8		0.04	
60 -					0.12	10
	75			0.13		
		44				
100 -			119			
- 150				0.11		25
150 -			171			50
- 250			188			
- 850	550	12	200			

Se pide:

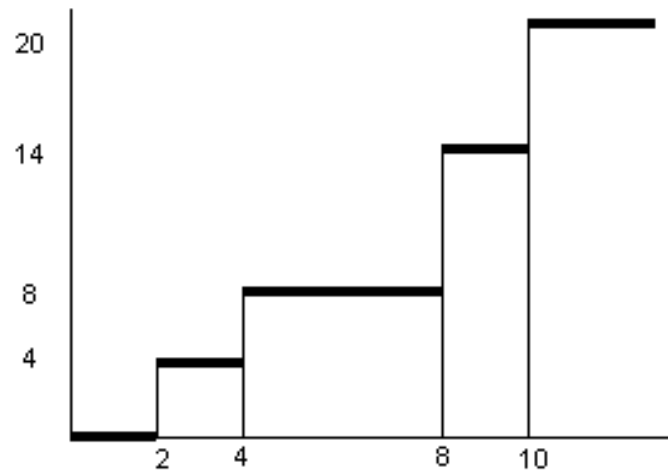
- Construir el histograma de frecuencias absolutas.
  - Calcular la media, moda y mediana.
  - ¿Qué porcentaje de personas presenta un ahorro superior a 120.000 ptas?.
  - ¿Entre qué valores de ahorro está situado el 80% central de las personas estudiadas?.
6. La siguiente gráfica recoge la información de la distribución de rentas familiares (en miles de \$) en los hogares de una cierta ciudad americana.



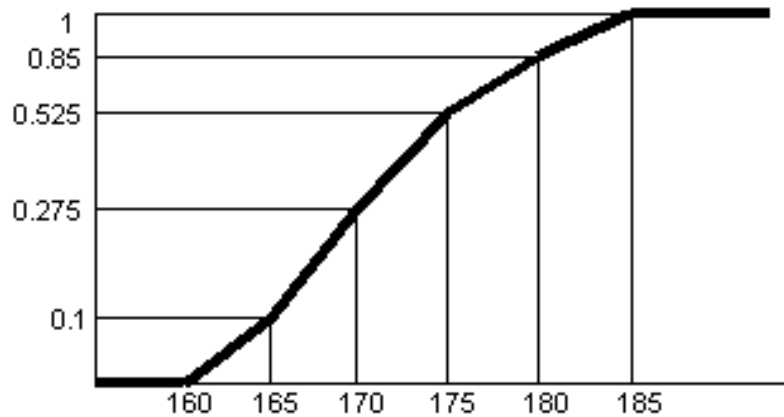
Se pide:

- ¿Qué tipo de variable se está analizando?. Construir la tabla de frecuencias asociada con estos datos.

- (b) Calcular la moda, media, mediana y desviación típica.
  - (c) ¿Qué porcentaje de hogares tiene una renta superior a 43.500 \$?.
7. El siguiente gráfico corresponde con las puntuaciones obtenidas por un grupo de alumnos. Se pide:



- (a) Observando el gráfico decir qué tipo de variable es la que se ha medido y de cuántos individuos consta la muestra.
  - (b) Contruir la tabla de frecuencias.
  - (c) Representar gráficamente la distribución de frecuencias relativas.
  - (d) ¿Cual es la frecuencia relativa acumulada del valor 6 de la variable?.
8. La distribución de las alturas de 40 estudiantes, obtenida de una lista alfabética de una Universidad, viene dada mediante el siguiente gráfico. Se pide:



- (a) Construir la tabla de frecuencias asociada con la variable altura y correspondiente a los datos representados.

- (b) ¿Cuántos alumnos miden entre 1.72 y 1.80?
  - (c) ¿Qué porcentaje de alumnos miden menos de 1.67?
  - (d) Dibujar el histograma de frecuencias de la variable en estudio.
  - (e) Encontrar el octavo decil y el coeficiente de variación de Pearson.
9. Se toma una muestra de llegadas de aviones a un aeropuerto para estudiar el problema de los retrasos en los vuelos. La siguiente tabla recoge el número de minutos de retraso en una muestra de 100 vuelos.

Min. Retraso	(0-10]	(10-20]	(20-30]	(30-40]	(40-50]	(50-60]
Número de Vuelos	29	23	17	14	11	6

Se pide:

- (a) Calcular el número de vuelos que sufren un retraso superior a 30 minutos.
  - (b) Calcular el número de vuelos que sufren un retraso inferior a 20 minutos.
  - (c) Calcular el número de vuelos cuyo retraso está comprendido entre 5 y 27 minutos.
  - (d) Si analizamos uno de los vuelos de la muestra y vemos que ha sufrido un retraso de 32 minutos, podríamos afirmar que el 50% de los vuelos de esta muestra han sufrido mayor retraso que éste. ¿Qué porcentaje de la muestra tiene un retraso inferior a los 32 minutos?
  - (e) ¿La variable retraso de las llegadas es simétrica?. ¿Cuál es la variación de los retrasos de las llegadas?.
10. Consideremos el siguiente par de muestras:

Muestra 1: 10, 9, 8, 7, 8, 6, 10, 6,  
Muestra 2: 10, 6, 10, 6, 8, 10, 8, 6,

Se pide:

- (a) Calcular el recorrido de ambas muestras. ¿Es posible concluir que ambas muestras exhiben la misma variabilidad?.
- (b) Calcular la desviación típica de cada una de las muestras. ¿Estas cantidades indican que ambas muestras tienen la misma variabilidad?.